

## 11.3 Collecting Data

**Essential Question** What are some considerations when undertaking a statistical study?

The goal of any statistical study is to collect data and then use the data to make a decision. Any decision you make using the results of a statistical study is only as reliable as the process used to obtain the data. If the process is flawed, then the resulting decision is questionable.

### EXPLORATION 1 Analyzing Sampling Techniques

**Work with a partner.** Determine whether each sample is representative of the population. Explain your reasoning.

- To determine the number of hours people exercise during a week, researchers use random-digit dialing and call 1500 people.
- To determine how many text messages high school students send in a week, researchers post a survey on a website and receive 750 responses.
- To determine how much money college students spend on clothes each semester, a researcher surveys 450 college students as they leave the university library.
- To determine the quality of service customers receive, an airline sends an e-mail survey to each customer after the completion of a flight.

### EXPLORATION 2 Analyzing Survey Questions

**Work with a partner.** Determine whether each survey question is biased. Explain your reasoning. If so, suggest an unbiased rewording of the question.

- Does eating nutritious, whole-grain foods improve your health?
- Do you ever attempt the dangerous activity of texting while driving?
- How many hours do you sleep each night?
- How can the mayor of your city improve his or her public image?

### EXPLORATION 3 Analyzing Survey Randomness and Truthfulness

**Work with a partner.** Discuss each potential problem in obtaining a random survey of a population. Include suggestions for overcoming the problem.

- The people selected might not be a random sample of the population.
- The people selected might not be willing to participate in the survey.
- The people selected might not be truthful when answering the question.
- The people selected might not understand the survey question.

## Communicate Your Answer

- What are some considerations when undertaking a statistical study?
- Find a real-life example of a biased survey question. Then suggest an unbiased rewording of the question.

### JUSTIFYING CONCLUSIONS

To be proficient in math, you need to justify your conclusions and communicate them to others.

# 11.3 Lesson

## Core Vocabulary

- random sample, p. 610
- self-selected sample, p. 610
- systematic sample, p. 610
- stratified sample, p. 610
- cluster sample, p. 610
- convenience sample, p. 610
- bias, p. 611
- unbiased sample, p. 611
- biased sample, p. 611
- experiment, p. 612
- observational study, p. 612
- survey, p. 612
- simulation, p. 612
- biased question, p. 613

### Previous

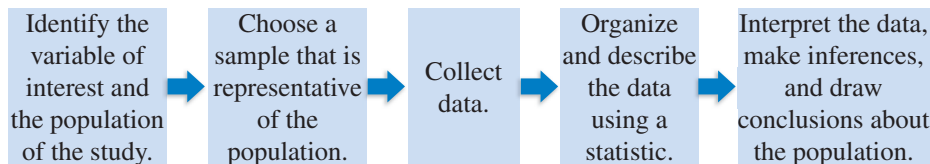
population  
sample

## What You Will Learn

- ▶ Identify types of sampling methods in statistical studies.
- ▶ Recognize bias in sampling.
- ▶ Analyze methods of collecting data.
- ▶ Recognize bias in survey questions.

## Identifying Sampling Methods in Statistical Studies

The steps in a typical statistical study are shown below.



There are many different ways of sampling a population, but a *random sample* is preferred because it is most likely to be representative of a population. In a **random sample**, each member of a population has an equal chance of being selected.

The other types of samples given below are defined by the methods used to select members. Each sampling method has its advantages and disadvantages.

## Core Concept

### Types of Samples

For a **self-selected sample**, members of a population can volunteer to be in the sample.



For a **systematic sample**, a rule is used to select members of a population. For instance, selecting every other person.



For a **stratified sample**, a population is divided into smaller groups that share a similar characteristic. A sample is then randomly selected from each group.



For a **cluster sample**, a population is divided into groups, called *clusters*. All of the members in one or more of the clusters are selected.



For a **convenience sample**, only members of a population who are easy to reach are selected.

### STUDY TIP

A stratified sample ensures that every segment of a population is represented.

### STUDY TIP

With cluster sampling, a member of a population cannot belong to more than one cluster.



### EXAMPLE 1 Identifying Types of Samples

You want to determine whether students in your school like the new design of the school's website. Identify the type of sample described.

- You list all of the students alphabetically and choose every sixth student.
- You mail questionnaires and use only the questionnaires that are returned.
- You ask all of the students in your algebra class.
- You randomly select two students from each classroom.

#### SOLUTION

- You are using a rule to select students. So, the sample is a *systematic* sample.
- The students can choose whether to respond. So, the sample is a *self-selected* sample.
- You are selecting students who are readily available. So, the sample is a *convenience* sample.
- The students are divided into similar groups by their classrooms, and two students are selected at random from each group. So, the sample is a *stratified* sample.

### Monitoring Progress Help in English and Spanish at [BigIdeasMath.com](http://BigIdeasMath.com)

- WHAT IF?** In Example 1, you divide the students in your school according to their zip codes, then select all of the students that live in one zip code. What type of sample are you using?
- Describe another method you can use to obtain a stratified sample in Example 1.

#### STUDY TIP

All good sampling methods rely on random sampling.

### Recognizing Bias in Sampling

A **bias** is an error that results in a misrepresentation of a population. In order to obtain reliable information and draw accurate conclusions about a population, it is important to select an *unbiased sample*. An **unbiased sample** is representative of the population that you want information about. A sample that overrepresents or under-represents part of the population is a **biased sample**. When a sample is biased, the data are invalid. A *random sample* can help reduce the possibility of a biased sample.

### EXAMPLE 2 Identifying Bias in Samples

Identify the type of sample and explain why the sample is biased.

- A news organization asks its viewers to participate in an online poll about bullying.
- A computer science teacher wants to know how students at a school most often access the Internet. The teacher asks students in one of the computer science classes.

#### SOLUTION

- The viewers can choose whether to participate in the poll. So, the sample is a *self-selected* sample. The sample is biased because people who go online and respond to the poll most likely have a strong opinion on the subject of bullying.
- The teacher selects students who are readily available. So, the sample is a *convenience* sample. The sample is biased because other students in the school do not have an opportunity to be chosen.

### EXAMPLE 3 Selecting an Unbiased Sample

You are a member of your school's yearbook committee. You want to poll members of the senior class to find out what the theme of the yearbook should be. There are 246 students in the senior class. Describe a method for selecting a random sample of 50 seniors to poll.

#### SOLUTION

- Step 1** Make a list of all 246 seniors. Assign each senior a different integer from 1 to 246.
- Step 2** Generate 50 unique random integers from 1 to 246 using the *randInt* feature of a graphing calculator.
- Step 3** Choose the 50 students who correspond to the 50 integers you generated in Step 2.

```
randInt(1,246)
84
245
50
197
235
55
```

#### STUDY TIP

When you obtain a duplicate integer during the generation, ignore it and generate a new, unique integer as a replacement.

#### Monitoring Progress Help in English and Spanish at [BigIdeasMath.com](http://BigIdeasMath.com)

- The manager of a concert hall wants to know how often people in the community attend concerts. The manager asks 45 people standing in line for a rock concert how many concerts they attend per year. Identify the type of sample the manager is using and explain why the sample is biased.
- In Example 3, what is another method you can use to generate a random sample of 50 students? Explain why your sampling method is random.

### Analyzing Methods of Data Collection

There are several ways to collect data for a statistical study. The objective of the study often dictates the best method for collecting the data.

#### Core Concept

##### Methods of Collecting Data

An **experiment** imposes a treatment on individuals in order to collect data on their response to the treatment. The treatment may be a medical treatment, or it can be any action that might affect a variable in the experiment, such as adding methanol to gasoline and then measuring its effect on fuel efficiency.

An **observational study** observes individuals and measures variables without controlling the individuals or their environment. This type of study is used when it is difficult to control or isolate the variable being studied, or when it may be unethical to subject people to a certain treatment or to withhold it from them.

A **survey** is an investigation of one or more characteristics of a population. In a survey, every member of a sample is asked one or more questions.

A **simulation** uses a model to reproduce the conditions of a situation or process so that the simulated outcomes closely match the real-world outcomes. Simulations allow you to study situations that are impractical or dangerous to create in real life.

#### READING

A *census* is a survey that obtains data from every member of a population. Often, a census is not practical because of its cost or the time required to gather the data. The U.S. population census is conducted every 10 years.

**EXAMPLE 4****Identifying Methods of Data Collection**

Identify the method of data collection each situation describes.

- a. A researcher records whether people at a gas station use hand sanitizer.
- b. A landscaper fertilizes 20 lawns with a regular fertilizer mix and 20 lawns with a new organic fertilizer. The landscaper then compares the lawns after 10 weeks and determines which fertilizer is better.

**SOLUTION**

- a. The researcher is gathering data without controlling the individuals or applying a treatment. So, this situation is an *observational study*.
- b. A treatment (organic fertilizer) is being applied to some of the individuals (lawns) in the study. So, this situation is an *experiment*.

**Monitoring Progress**

Help in English and Spanish at [BigIdeasMath.com](http://BigIdeasMath.com)

Identify the method of data collection the situation describes.

5. Members of a student council at your school ask every eighth student who enters the cafeteria whether they like the snacks in the school's vending machines.
6. A park ranger measures and records the heights of trees in a park as they grow.
7. A researcher uses a computer program to help determine how fast an influenza virus might spread within a city.

**STUDY TIP**

Bias may also be introduced in survey questioning in other ways, such as by the order in which questions are asked or by respondents giving answers they believe will please the questioner.

**Recognizing Bias in Survey Questions**

When designing a survey, it is important to word survey questions so they do not lead to biased results. Answers to poorly worded questions may not accurately reflect the opinions or actions of those being surveyed. Questions that are flawed in a way that leads to inaccurate results are called **biased questions**. Avoid questions that:

- encourage a particular response
- do not provide enough information to give an accurate opinion
- are too sensitive to answer truthfully
- address more than one issue

**EXAMPLE 5****Identify and Correct Bias in Survey Questioning**

A dentist surveys his patients by asking, “Do you brush your teeth at least twice per day and floss every day?” Explain why the question may be biased or otherwise introduce bias into the survey. Then describe a way to correct the flaw.

**SOLUTION**

Patients who brush less than twice per day or do not floss daily may be afraid to admit this because the dentist is asking the question. One improvement may be to have patients answer questions about dental hygiene on paper and then put the paper anonymously into a box.

**Monitoring Progress**

Help in English and Spanish at [BigIdeasMath.com](http://BigIdeasMath.com)

8. Explain why the survey question below may be biased or otherwise introduce bias into the survey. Then describe a way to correct the flaw.

“Do you agree that our school cafeteria should switch to a healthier menu?”

## Vocabulary and Core Concept Check

- VOCABULARY** Describe the difference between a stratified sample and a cluster sample.
- COMPLETE THE SENTENCE** A sample for which each member of a population has an equal chance of being selected is a(n) \_\_\_\_\_ sample.
- WRITING** Describe a situation in which you would use a simulation to collect data.
- WRITING** Describe the difference between an unbiased sample and a biased sample. Give one example of each.

## Monitoring Progress and Modeling with Mathematics

In Exercises 5–8, identify the type of sample described. (See Example 1.)

- The owners of a chain of 260 retail stores want to assess employee job satisfaction. Employees from 12 stores near the headquarters are surveyed.
- Each employee in a company writes their name on a card and places it in a hat. The employees whose names are on the first two cards drawn each win a gift card.
- A taxicab company wants to know whether its customers are satisfied with the service. Drivers survey every tenth customer during the day.
- The owner of a community pool wants to ask patrons whether they think the water should be colder. Patrons are divided into four age groups, and a sample is randomly surveyed from each age group.

In Exercises 9–12, identify the type of sample and explain why the sample is biased. (See Example 2.)

- A town council wants to know whether residents support having an off-leash area for dogs in the town park. Eighty dog owners are surveyed at the park.



- A sportswriter wants to determine whether baseball coaches think wooden bats should be mandatory in collegiate baseball. The sportswriter mails surveys to all collegiate coaches and uses the surveys that are returned.
- You want to find out whether booth holders at a convention were pleased with their booth locations. You divide the convention center into six sections and survey every booth holder in the fifth section.
- Every tenth employee who arrives at a company health fair answers a survey that asks for opinions about new health-related programs.
- ERROR ANALYSIS** Surveys are mailed to every other household in a neighborhood. Each survey that is returned is used. Describe and correct the error in identifying the type of sample that is used.



Because the surveys were mailed to every other household, the sample is a systematic sample.

- ERROR ANALYSIS** A researcher wants to know whether the U.S. workforce supports raising the minimum wage. Fifty high school students chosen at random are surveyed. Describe and correct the error in determining whether the sample is biased.



Because the students were chosen at random, the sample is not biased.

**In Exercises 15–18, determine whether the sample is biased. Explain your reasoning.**

15. Every third person who enters an athletic event is asked whether he or she supports the use of instant replay in officiating the event.
16. A governor wants to know whether voters in the state support building a highway that will pass through a state forest. Business owners in a town near the proposed highway are randomly surveyed.
17. To assess customers' experiences making purchases online, a rating company e-mails purchasers and asks that they click on a link and complete a survey.
18. Your school principal randomly selects five students from each grade to complete a survey about classroom participation.

19. **WRITING** The staff of a student newsletter wants to conduct a survey of the students' favorite television shows. There are 1225 students in the school. Describe a method for selecting a random sample of 250 students to survey. (See Example 3.)



20. **WRITING** A national collegiate athletic association wants to survey 15 of the 120 head football coaches in a division about a proposed rules change. Describe a method for selecting a random sample of coaches to survey.

**In Exercises 21–24, identify the method of data collection the situation describes. (See Example 4.)**

21. A researcher uses technology to estimate the damage that will be done if a volcano erupts.



22. The owner of a restaurant asks 20 customers whether they are satisfied with the quality of their meals.
23. A researcher compares incomes of people who live in rural areas with those who live in large urban areas.
24. A researcher places bacteria samples in two different climates. The researcher then measures the bacteria growth in each sample after 3 days.

**In Exercises 25–28, explain why the survey question may be biased or otherwise introduce bias into the survey. Then describe a way to correct the flaw. (See Example 5.)**

25. “Do you agree that the budget of our city should be cut?”
26. “Would you rather watch the latest award-winning movie or just read some book?”
27. “The tap water coming from our western water supply contains twice the level of arsenic of water from our eastern supply. Do you think the government should address this health problem?”
28. A child asks, “Do you support the construction of a new children’s hospital?”

**In Exercises 29–32, determine whether the survey question may be biased or otherwise introduce bias into the survey. Explain your reasoning.**

29. “Do you favor government funding to help prevent acid rain?”
30. “Do you think that renovating the old town hall would be a mistake?”
31. A police officer asks mall visitors, “Do you wear your seat belt regularly?”
32. “Do you agree with the amendments to the Clean Air Act?”

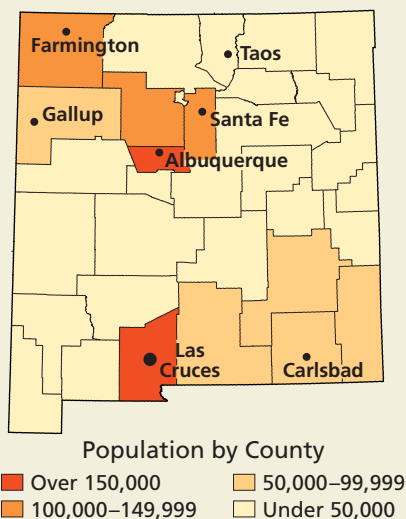
33. **REASONING** A researcher studies the effect of fiber supplements on heart disease. The researcher identified 175 people who take fiber supplements and 175 people who do not take fiber supplements. The study found that those who took the supplements had 19.6% fewer heart attacks. The researcher concludes that taking fiber supplements reduces the chance of heart attacks.



- a. Explain why the researcher’s conclusion may not be valid.
- b. Describe how the researcher could have conducted the study differently to produce valid results.

**34. HOW DO YOU SEE IT?** A poll is conducted to predict the results of a statewide election in New Mexico before all the votes are counted. Fifty voters in each of the state's 33 counties are asked how they voted as they leave the polls.

- Identify the type of sample described.
- Explain how the diagram shows that the polling method could result in a biased sample.



**35. WRITING** Consider each type of sample listed on page 610. Which of the samples are most likely to lead to biased results? Explain.

**36. THOUGHT PROVOKING** What is the difference between a “blind experiment” and a “double-blind experiment?” Describe a possible advantage of the second type of experiment over the first.

**37. WRITING** A college wants to survey its graduating seniors to find out how many have already found jobs in their field of study after graduation.

- What is the objective of the survey?
- Describe the population for the survey.
- Write two unbiased questions for the survey.

**38. REASONING** About 3.2% of U.S. adults follow a vegetarian-based diet. Two randomly selected groups of people were asked whether they follow such a diet. The first sample consists of 20 people and the second sample consists of 200 people. Which sample proportion is more likely to be representative of the national percentage? Explain.

**39. MAKING AN ARGUMENT** The U.S. Census is taken every 10 years to gather data from the population. Your friend claims that the sample cannot be biased. Is your friend correct? Explain.

**40. OPEN-ENDED** An airline wants to know whether travelers have enough leg room on its planes.



- What method of data collection is appropriate for this situation?
- Describe a sampling method that is likely to give biased results. Explain.
- Describe a sampling method that is *not* likely to give biased results. Explain.
- Write one biased question and one unbiased question for this situation.

**41. REASONING** A website contains a link to a survey that asks how much time each person spends on the Internet each week.

- What type of sampling method is used in this situation?
- Which population is likely to respond to the survey? What can you conclude?

## Maintaining Mathematical Proficiency

Reviewing what you learned in previous grades and lessons

Evaluate the expression without using a calculator. (Section 5.1)

42.  $4^{5/2}$

43.  $27^{2/3}$

44.  $-64^{1/3}$

45.  $8^{-2/3}$

Simplify the expression. (Section 5.2)

46.  $(4^{3/2} \cdot 4^{1/4})^4$

47.  $(6^{1/3} \cdot 3^{1/3})^{-2}$

48.  $\sqrt[3]{4} \cdot \sqrt[3]{16}$

49.  $\frac{\sqrt[4]{405}}{\sqrt[4]{5}}$